MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190319

A Decision Support System for Pathology Test Result Reviews in an Emergency Department to Support Patient Safety and Increase Efficiency

Anthony Nguyen[a], Hamed Hassanzadeh[a], Yushi Zhang[a], John O'Dwyer[a], David Conlan[a], Michael Lawley[a], Jim Steel[a], Kylynn Loi[a], Peter Rizzo[b]

[a] The Australian e-Health Research Centre, CSIRO, Brisbane, Queensland, Australia,
[b] The Prince Charles Hospital, Department of Health, Brisbane, Queensland, Australia

Abstract

The review of pathology test results for missed diagnoses in Emergency Departments is time-consuming, laborious, and can be inaccurate. An automated solution, with text mining and clinical terminology semantic capabilities, was developed to provide clinical decision support. The system focused on the review of microbiology test results that contained information on culture strains and their antibiotic sensitivities, both of which can have a significant impact on ongoing patient safety and clinical care. The system was highly effective at identifying abnormal test results, reducing the number of test results for review by 92%. Furthermore, the system reconciled antibiotic sensitivities with documented antibiotic prescriptions in discharge summaries to identify patient follow-ups with a 91% F-measure – allowing for the accurate prioritization of cases for review. The system dramatically increases accuracy, efficiency, and supports patient safety by ensuring important diagnoses are recognized and correct antibiotics are prescribed.

Keywords:

Clinical Decision Support Systems, Data Mining, Emergency Medicine

Introduction

The failure to follow-up on pending test results when patients have been discharged from hospitals is a potential threat to patient safety [1][3]. Poor test result follow-up can significantly impact patient safety and clinical care, including missed diagnoses and suboptimal patient outcomes [3]. The Emergency Care Research Institute (ECRI) underscored the problem of "Test result reporting and follow-up" as one of the top 10 patient safety concerns in healthcare organizations [4].

This is especially problematic in Emergency Departments (EDs) where there can be a significant number of pending test results when patients have been discharged from the hospital. The current manual process in EDs for checking test results for abnormalities, then ensuring patients with abnormal results are appropriately followed-up is sub-optimal [5][6]. The process is labor-intensive, un-prioritized by its nature, and has negative impacts on patient care and staff workloads. Specifically, the current process 1) potentially delays the follow-up of critical abnormal cases because all test results need to be reviewed then reconciled against the patient's disposition recorded in the ED; 2) consumes valuable clinical hours that could be better spent engaged in direct patient care; and 3) is prone to errors due to the time intensive nature of work trying to find the "needle in the haystack" of results, in addition to ongoing pressures of clinician workloads.

Furthermore, the process of checking test results in EDs has not kept pace with the rapid growth and expansion of hospital services due to increasing population, aging, and chronic diseases. This has placed increased pressures on "systems" and staff workloads, leading to reduced efficiencies and reduced direct clinical contact hours with patients potentially compromising patient safety and care.

Proposed is an automated solution that changes the way EDs detect and report abnormal test results. The system initially focused on the review of microbiology test results as these can take up to several days to be reported upon, by which time a patient will have been discharged from the ED. These test results often contain information on the results of culture strains present and their antibiotic sensitivities, both of which make a significant impact on ongoing patient safety and clinical care.

In particular, contributions include 1) 'trigger' algorithm to identify abnormal microbiology test results and related antibiotic sensitivities from pathology reports; 2) extraction of antibiotic prescriptions documented in ED discharge summaries using text mining based 'trigger' keywords derived from clinical terminology semantics based on SNOMED CT[1] to reconcile against the antibiotic sensitivities; and 3) automate a protocol with a prioritized listing of cases to support the test result follow-up checking process. The solution overcomes the labor-intensive and error-prone nature of test result reviews to enhance patient safety and efficiency.

Background

Clinical decision support (CDS) systems have the potential to improve clinical efficiency and patient safety [7][11]. These systems have been applied across diverse settings and range from information management systems through to clinical alerts, and diagnosis and/or treatment recommendations. Despite the significant body of literature on CDS systems, automation in the context of test result reviewing has been limited.

Health system processes and test result management systems have been implemented with the aim to improve test result review processes [3]. However, physicians are still unsatisfied with how they manage test results [5][6]. For example, an end-to-end workflow is much desired along with the ability to filter normal (irrelevant) results to help prioritize the workflow. This

[1] Systematized nomenclature of medicine – clinical terms

suggests that current systems do not include robust CDS features or functions such as prioritization and/or filtering to the end user [8].

Automated notifications of abnormal test results can significantly reduce the number of results for review. Abnormality detection from pathology and imaging test results range from 'trigger' algorithms that use clinical logic on structured electronic health record (EHR) data to identify abnormalities [12][15] through to advanced computational approaches such as machine learning to identify abnormalities from unstructured narrative EHR data [16][18].

Furthermore, the abnormality task can be extended to also reconcile or link the abnormal findings in radiology reports with the patient's disposition recorded in ED information system to provide decision support to the manual review process [16].

Here, the proposed solution for microbiology test result reviews leverages previous work on 'trigger' algorithms and reconciliation to develop an end-to-end test result reconciliation solution. Microbiology test results most frequently report on bacterial antimicrobial (or antibiotic) susceptibility. Antibiotics have been used to treat infections or diseases caused by bacteria and have saved many lives since their introduction. However, antibiotics may not be effective when overused due to antimicrobial resistance. This "antimicrobial stewardship" problem is a global problem and like "Test result reporting and follow-up," it was also identified as being within the top 10 patient safety concerns in healthcare organizations [4]. This end-to-end solution can aid in ensuring that appropriate antibiotic prescription continues through the treatment of the infection thus fulfilling the needs for antimicrobial stewardship.

Methods

The workflow for identification and reconciliation of abnormal microbiology test results with ED discharge summaries will be presented along with the proposed methodology for automating it. The workflow was adapted from an actual test result review process within an ED in Brisbane, Australia.

Data

The dataset was obtained from The Prince Charles Hospital (TPCH), Brisbane, Australia[2]. The dataset comprised 31,787 ED encounters (pertaining to 15,916 unique patients) from July 2013 to December 2014 (18-month period). A separate dataset of microbiology test results, ordered from the same hospital and ED, was obtained from a state-wide pathology information system and comprised 29,503 microbiology pathology HL7 messages (pertaining to 18,560 patients) from a wider 4-year time period. The ED encounters and microbiology test results were matched based on their unique patient identifier and test result order date is within the patient's ED admission and discharge date/time. A total of 16,867 ED encounters had matching microbiology test results over the dataset time period.

A subset of abnormal test results and matching ED discharge summaries (142 cases) was manually reviewed by ED senior medical officers to determine their follow-up requirement. Eighty percent of the cases (113 cases) were used for system development while the remaining cases were used for testing (29 cases). An additional 282 cases were subsequently obtained to assess the generalisability and robustness of the system via a pilot study. Table 1 presents a summary of the gold standard dataset.

Table 1 – Gold standard dataset for abnormal microbiology test result reconciliation

Dataset	Follow-up	No follow-up
Development set	78 (69%)	35 (31%)
Test set	20 (69%)	9 (31%)
Pilot set	211 (75%)	71 (25%)

Abnormal test result identification and filtering

The TPCH ED, on a daily basis (including weekends), would print out the microbiology test results and place them in a dedicated area for sorting. Results must be sorted by hand to identify abnormal results requiring follow-up from those that do not. This process was not without errors. A typical day may find the sorting pile overlooked or half completed due to clinical demands. Other errors include failure to recognize abnormal results from distraction or incorrect interpretation of results.

The proposed system addressed the abnormal test result identification issue by utilizing the concept of the trigger algorithm [12][15]. Triggers were applied to the microbiology test results to identify abnormal results and related antibiotic sensitivities. The presence or absence of antibiotic sensitivity results in pathology HL7 messages was used as the trigger.

The set of abnormal test results could be further filtered based on the patient's discharge destination. Patients who were admitted into the hospital as an in-patient would not be required to be followed-up as they would be appropriately followed-up by other clinicians in the hospital. To achieve this, the system again applied triggers to the discharge destination field in the ED information system to identify patients who were not admitted to any of the hospital wards. A list of all possible ED discharge destinations with their follow-up requirement was provided for the discharge destination filtering.

The processing and filtering of pathology reports resulted in a significant reduction in the number of test results requiring review by the ED.

Abnormal test result reconciliation

The next stage of the review process was abnormal test result reconciliation whereby abnormal test results must be correlated against the clinical record. The ED clinician searches for patients with abnormal test results, one by one, in their ED medical record to determine if patients were required to be followed-up due to an inappropriate diagnosis or treatment. This correlation process was also not without errors. For example, appropriate action may not have been taken due to failure to recognize the misdiagnosis or incorrect treatment.

This stage involved the application of text mining and clinical terminology semantics for the extraction of antibiotic prescriptions documented in ED discharge summaries for reconciling against antibiotic sensitivities in pathology test results. To extract antibiotic prescriptions in discharge summaries, a 'trigger' keyword list containing antibiotic names was compiled. This was used to match occurrences of these keywords in discharge summaries.

A baseline list of antibiotic 'trigger' keywords was derived from the list of possible antibiotic sensitivities identified in microbiology test results (e.g. Trimethoprim, Co-trimoxazole,

[2] Research ethics was obtained from the Metro North Hospital and Health Services Human Research Ethics Committee.

and Di(Flu)cloxacillin). Antibiotic names and their expanded forms (e.g. dicloxacillin and flucloxacillin for Di(Flu)cloxacillin) would form candidate antibiotics to use as 'trigger' keywords.

An extended 'trigger' keyword list was compiled by supplementing the antibiotic names with also their brand names. Here, the Australian Medicines Terminology (AMT)[3] and the SNOMED CT expression constraint language (ECL)[4] was used to generate the list of trade names. Ontoserver [19], a clinical terminology server with support for SNOMED CT and AMT as well as SNOMED CT's ECL, was used to generate the list of antibiotic trade names.

ECL templates were devised to return a list of trade names given 1) a single active antibiotic ingredient, and 2) more than one active antibiotic ingredients. An example ECL for a single antibiotic ingredient `2691011000036102|Trimethoprim|` is as follows:

```
^ 929360021000036102|Trade product reference set|
AND >> (
  ( ^ 929360031000036100|Trade product unit of use
reference set| : (
      700000081000036101|has intended active
ingredient| = 2691011000036102|Trimethoprim|
      ) AND [1..1] 700000081000036101|has intended
active ingredient| = * )
)
```

Ontoserver provides an ECL high-level reference[5] to aid in the interpretation of the above expression. In brief, the ECL returns a list of trade names (from `Trade product reference set`) that contains medications (from `Trade product unit of use reference set`) that have the specified antibiotic as its active ingredient. The `[1..1]` constrains the results to medications with only a single ingredient.

If multiple antibiotic ingredients were applicable to a certain antibiotic sensitivity test, then the above ECL can be adapted to include additional antibiotics using the `OR` operator. An example ECL extract for 'Di(Flu)cloxacillin' would be as follows:

```
...

700000081000036101|has intended active ingredient| =
  ( 2018011000036104|dicloxacillin| OR
    2115011000036102|flucloxacillin| )

...
```

For cases where an antibiotic has more than one active ingredient such as 'Co-trimoxazole' which contains both `2605011000036103|sulfamethoxazole|` and `2691011000036102|trimethoprim|`, then the following ECL was applied:

```
^ 929360021000036102|Trade product reference set|
AND >> (
  ( ^ 929360031000036100|Trade product unit of use
reference set| : (
      700000081000036101|has intended active
ingredient| = 2605011000036103|sulfamethoxazole|,
      700000081000036101|has intended active
ingredient| = 2691011000036102|trimethoprim|) )
)
```

The resulting list of antibiotic 'trigger' keywords was used to extract antibiotics documented in discharge summaries.

Discharge summaries often document the antibiotic treatments administered during the ED encounter. Intravenous (IV) medications for certain antibiotics can only be administered as IV-only. These IV-only antibiotics are generally not the full course of antibiotics and thus not relevant for reconciling against antibiotic sensitivities in test results. A list of all IV-only antibiotics was provided for antibiotic filtering. The resultant non-IV-only antibiotics would be used for the next stage of reconciliation.

The reconciliation of antibiotic prescriptions extracted from discharge summaries against antibiotic sensitivities from test results applied the following rule-based logic. The patient would not be followed-up if they were prescribed with an antibiotic that had a corresponding antibiotic sensitivity in their test result of 'sensitive'. However, the patient would require follow-up if any of the following reconciliation events occur:

- The patient has not been prescribed any antibiotics.
- The prescribed antibiotic was not tested (and thus its antibiotic sensitivity was unknown).
- The prescribed antibiotic resulted in an antibiotic sensitivity of 'resistant.'

The logic would be applied to each culture strain (or bacterial organism) identified in the test result.

Evaluation Measures

The efficiency of the system was evaluated based on the resultant number of test results identified by the system for clinical review (normalized on a weekly basis) compared to the full set of test results that would have been manually reviewed.

The reconciliation phase to determine whether or not a patient required follow-up was evaluated against the gold standard. The effectiveness of the system was measured using sensitivity (or recall) and positive predictive value (PPV or precision). To provide a single, overall evaluation measure, precision and recall were combined into a third evaluation measure, F-measure.

Results

A total of 16,867 ED encounters had matching microbiology test results over the dataset time period. This averages to ~216 test results per week, which an ED clinician would need to sort and review the test results.

Abnormal test result identification and filtering

The efficiency of the system in filtering irrelevant test results is tabulated in Table 2.

The proposed trigger to filter normal (irrelevant) test results from abnormal test results resulted in 2,605 abnormal test results – an average of 33 reports per week requiring review. The trigger was confirmed by ED clinicians to be accurate in identifying abnormal from normal test results.

[3] AMT is a subset of SNOMED CT-AU (Australian extension) for medicines commonly used in Australia.

[4] ECL is a formal language for defining bounded sets of clinical meanings represented by pre-coordinated or post-coordinated expressions.

[5] https://ontoserver.csiro.au/shrimp/ecl_help.html

Table 2 – System efficiency results in terms of reducing the number of abnormal test results review

	Number of test results	Weekly number of test results
Full manual review	16,867	216
+ Abnormality identification	2,605	33 (↓ 85%)
+ Discharge destination filtering	1,379	18 (↓ 92%)

When the abnormal test results were filtered based on the ED discharge destination, the number of test results that actually required clinical review was reduced to 1,379 (18 reports per week).

Abnormal test result reconciliation

To further provide clinical decision support to the test result review process, the reconciliation of antibiotic prescriptions extracted from discharge summaries against antibiotic sensitivities in test results allows for the prioritization of cases for clinical review. Table 3 presents the classification effectiveness of the proposed reconciliation approach.

Table 3 – System effectiveness results in classifying the follow-up requirement

Dataset	PPV	Sensitivity	F-measure
Development set	0.802	0.936	0.869
Test set	0.905	0.950	0.900
Pilot set	0.858	0.943	0.898

Discussion

The automated identification and prioritization of abnormal microbiological pathology reports for clinical review will bring about benefits in cost efficiencies, quality of care, and performance.

Noteworthy, was that a simple solution based on triggers was able to substantially reduce the number test results for review. Results show a 92% reduction in microbiology test results that required review.

Furthermore, the system was able to accurately identify the follow-up requirements (F-measure of 90%) for the prioritization of cases for clinical review. The very limited differences in performances between the development, test and pilot dataset show the generalisability and robustness of the proposed antibiotic extraction approach using 'trigger' keywords and a rule-based reconciliation logic.

Error analysis on the development dataset revealed four error categories: 1) antibiotic misspellings in discharge summaries (4 cases); 2) context of antibiotic mentions (3 cases; e.g., previous prescriptions and non-IV-only administered antibiotics); 3) extrapolation of prescribed antibiotics with sensitivity results (4 cases; e.g., cephalexin prescription considered a correct prescription for an organism sensitive to Cefazolin test result); 4) missed identification of antibiotic mentions in discharge summaries (1 case); and 5) gold standard inaccuracies confirmed to be human errors (7 cases). These error categories form avenues for future work to improve system performances.

Conclusions

The proposed IT solution combines text mining and decision support technologies for the novel identification and reconciliation of abnormal microbiology test results in EDs. It has the potential to dramatically increase the accuracy and efficiency of microbiology test result review to support patient safety by ensuring important diagnoses are recognized and correct antibiotics have been prescribed. The increased efficiency will allow significantly more clinical hours devoted to the direct treatment of patients presenting to hospital EDs to increase their quality of care.

Next steps include the planning and development of end-user software that would allow for appropriate presentation without impeding on clinical workflow [7][20]. The clinical decision support system would then be implemented and trialed by clinicians in actual practice.

The proposed 'triggers' and reconciliation logic is applicable to any software system using HL7 communications. This would allow changes to be made among different pathology systems or EHRs to provide the notification and reconciliation of abnormal test results. It can also be applied to other areas of pathology as well as the reporting of radiology test results, which are similar in processes.

Acknowledgements

We thank Dr. Mike Watson of TPCH ED for assisting with the initial project management and data extraction; Sally Taylor of TPCH Pharmacy Department, Dr Robert Horvath, Dr Alex Chaudhuri and Dr John McNamara of TPCH Infectious Disease Department for their domain expertise and advice; and Doreen Lelieveld for assisting Dr Peter Rizzo with the review of microbiology test results and discharge summaries for patient follow-ups.

References

[1] Roy CL, Poon EG, Karson AS, Ladak-Merchant Z, Johnson RE, Maviglia SM, Gandhi TK. Patient safety concerns arising from test results that return after hospital discharge. Ann Intern Med. 2005 Jul 19;143(2):121-8.

[2] Walz SE, Smith M, Cox E, Sattin J, Kind AJ. Pending laboratory tests and the hospital discharge summary in patients discharged to sub-acute care. J Gen Intern Med. 2011 Apr 1;26(4):393-8.

[3] Callen J, Georgiou A, Li J, Westbrook JI. The impact for patient outcomes of failure to follow up on test results. How can we do better? EJIFCC. 2015 Jan;26(1):38.

[4] ECRI Institute. Top 10 patient safety concerns for healthcare organizations. 2017.

[5] Poon EG, Gandhi TK, Sequist TD, Murff HJ, Karson AS, Bates DW. "I wish I had seen this test result earlier!": dissatisfaction with test result management systems in primary care. Arch Intern Med. 2004 Nov 8;164(20):2223-8.

[6] Dalal AK, Poon EG, Karson AS, Gandhi TK, Roy CL. Lessons learned from implementation of a computerized application for pending tests at hospital discharge. J Hosp Med. 2011 Jan;6(1):16-21.

[7] Musen MA, Middleton B, Greenes RA. Clinical decision-support systems. In Biomedical informatics 2014 (pp. 643-674). Springer, London.

[8] Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, Campbell E, Bates DW. Grand challenges in clinical decision support. J Biomed Inf. 2008 Apr 1;41(2):387-92.

[9] Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, Samsa G, Hasselblad V, Williams JW, Musty MD, Wing L. Effect of clinical decision-support systems: a systematic review. Ann. Intern. Med. 2012 Jul 3;157(1):29-43.

[10] Miller A, Moon B, Anders S, Walden R, Brown S, Montella D. Integrating computerized clinical decision support systems into clinical work: a meta-synthesis of qualitative research. Int J Med Inform. 2015 Dec 1;84(12):1009-18.

[11] Demner-Fushman D, Chapman WW, McDonald CJ. What can natural language processing do for clinical decision support? J Biomed Inf. 2009 Oct 1;42(5):760-72.

[12] Rosen AK, Mull HJ, Kaafarani H, Nebeker J, Shimada S, Helwig A, Nordberg B, Long B, Savitz LA, Shanahan CW, Itani K. Applying trigger tools to detect adverse events associated with outpatient surgery. J Patient Saf. 2011 Mar 1;7(1):45-59.

[13] Murphy DR, Laxmisan A, Reis BA, Thomas EJ, Esquivel A, Forjuoh SN, Parikh R, Khan MM, Singh H. Electronic health record-based triggers to detect potential delays in cancer diagnosis. BMJ Qual Saf. 2014 Jan 1;23(1):8-16.

[14] Murphy DR, Meyer AN, Bhise V, Russo E, Sittig DF, Wei L, Wu L, Singh H. Computerized triggers of big data to detect delays in follow-up of chest imaging results. Chest. 2016 Sep 1;150(3):613-20.

[15] Singh H, Thomas EJ, Sittig DF, Wilson L, Espadas D, Khan MM, Petersen LA. Notification of abnormal lab test results in an electronic medical record: do any safety concerns remain? Am J Med. 2010 Mar 1;123(3):238-44.

[16] Koopman B, Zuccon G, Wagholikar A, Chu K, O'Dwyer J, Nguyen A, Keijzers G. Automated reconciliation of radiology reports and discharge summaries. AMIA 2015 (Vol. 2015, pp. 775-784).

[17] Hassanzadeh H, Nguyen A, Karimi S, Chu K. Transferability of artificial neural networks for clinical document classification across hospitals: A case study on abnormality detection from radiology reports. J Biomed Inform. 2018 Sep 1;85:68-79.

[18] Hassanzadeh H, Kholghi M, Nguyen A, Chu K. Clinical Document Classification Using Labeled and Unlabeled Data Across Hospitals. AMIA 2018 (Vol. 2018, pp. 545-554.

[19] Metke-Jimenez A, Steel J, Hansen D, Lawley M. Ontoserver: a syndicated terminology server. J Biomed Semantics. 2018 Dec;9(1):24.

[20] Horsky J, Schiff GD, Johnston D, Mercincavage L, Bell D, Middleton B. Interface design principles for usable decision support: a targeted review of best practices for clinical prescribing interventions. J Biomed Inf. 2012 Dec 1;45(6):1202-16.

Address for correspondence

Anthony Nguyen, Email: Anthony.Nguyen@csiro.au.